THE DETERMINATION OF

EMPIRICAL AND ANALYTICAL SPACECRAFT PARAMETRIC CURVES

- THEORY AND METHODS -

PROGRESS REPORT III

Submitted by

TEXAS A&M RESEARCH FOUNDATION

to the

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

for

NASA GRANT SC-NGR-44-001-027

FACILITY FORM 602

N67-16114 (ACCESSION NUMBER) N67-16117 (THRU)

61 (PAGES) O (CODE)

CR-80921 (NASA CR OR TMX OR AD NUMBER) [redacted] 34 (CATEGORY)

Prepared by

INDUSTRIAL ENGINEERING DEPARTMENT

TEXAS A&M UNIVERSITY

November 1, 1966

N67-16115

PART I

DYNAMIC PROGRAMMING ALGORITHM FOR DETERMING 'BEST FIT'

By

Glen Self

For

Industrial Engineering Department

Texas A&M University

November 1966

FOREWORD

This document represents the third progress report on the NASA research grant NGR 44-001-027. The report is divided into four parts. A summary of these parts is presented below in order to outline the general nature of each section.

Part I is a revision of a section of a previous report which developed an algorithm for utilization of dynamic programming for determining the best fit or best combination of single estimating relationships into a composite estimating relationship. It has been extended to include an example to illustrate the mechanics of the algorithm.

Part II is a summary of a major part of the research conducted under this grant up to this point in time. It is an effort to bring the various parts of previously reported research into a proper perspective with each other.

Part III is an extension of work previously reported in Progress Report II. This section contains an explanation of the appended computer program and its use for computation of run-out costs for subsystems presently being developed.

Part IV is a completed area of research. This report contains the results of using expertise in the formulation of mathematical models. Due to its integral nature, it is submitted as a separate document.

DYNAMIC PROGRAMMING ALGORITHM FOR DETERMINING 'BEST FIT'

by

Glen Self
Texas A&M University

The problem of weighting individual predictors into a single predictor is one which can be approached by more than one method. One method which was reviewed recently was that of subjectively weighting the individual predictors which contained a single independent variable according to the importance that variable was felt to have on the function being considered. This particular approach did not produce consistently good results and did not present a quantitative base for making decisions as to whether the individual predictors were at fault or the weighting scheme being used. Therefore, by expanding the problem to (1) one of having any number of terms that were to be combined into a single function under the constraint that the sum of the weights used should approximate one (2) one of being further restricted by a small number of data points (which precipitated the original problem of only using a single variable predictor for a given type of function and still have some degrees of freedom associated with the error sum of squares) (3) one of having an inherent flexibility such that consecutive last terms could be deleted from consideration and still provide the optimum solution without recomputation, (4) providing for a built-in sensitivity whereby the effect of variation in the weightings could be evaluated and (5) where the minimum sum of squares criteria could be used as a basis for determining optimum weighting of the terms.

In the dynamic programming terminology the stages correspond to the terms which are being combined; therefore, it is necessary to develop the recursive relation of dynamic programming. For example if the actual values are y_i and the individual predicted values for the first term is x_{1i}, then the objective is to select some value θ_1 such that the function

$$f_1(\theta) = \sum_{i=1}^{n} (\theta_1 x_{1i} - y_i)^2 \quad \text{is}$$

minimized subject to $0 \leq \theta_1 \leq \theta$. Using this same notation scheme, the objective for a two-term equation is to determine

$$f_2(\theta) = \min_{0 \leq \theta_2 \leq \theta} \left[\sum_{i=1}^{n} (\theta_1 x_{1i} + \theta_2 x_{2i} - y_i)^2 \right]$$

$$= \min_{0 \leq \theta_2 \leq \theta} \left[\sum_{i=1}^{n} (\theta_1^2 x_{1i}^2 + \theta_2^2 x_{2i}^2 + y_i^2 + 2\theta_1 x_{1i} \theta_2 x_{2i} \right.$$

$$\left. - 2\theta_1 x_{1i} y_i - 2\theta_2 x_{2i} y_i \right]$$

However, $$f_1(\theta) = \min_{0 \leq \theta_1 \leq \theta} \left[\sum_{i=1}^{n} (\theta_1^2 x_{1i}^2 - 2\theta_1 x_{1i} + y_i^2) \right]$$

and $$\sum_{i=1}^{n} (\theta_2 x_{2i} - y_i)^2 = \sum_{i=1}^{n} (\theta_2^2 x_{2i}^2 - 2\theta_2 x_{2i} y_i + y_i^2)$$

$$\therefore f_2(\theta) = \min_{0 \leq \theta_2} \left[\sum_{i=1}^{n} (\theta_2 x_{2i} - y_i)^2 - \sum_{i=1}^{n} y_i^2 \right.$$

$$\left. + 2 \sum_{i=1}^{n} \theta_1 x_{1i} \theta_2 x_{2i} + f_1(\theta - \theta_2) \right].$$

It is probably only of interest at this point, but should be clarified for the next stage of the computation is that the values of $\theta_1 X_{1i}$ are fixed, based upon the value θ_1 takes on in order to optimize $f_1\ (\theta - \theta_2)$ i.e., some value of $0 \leq \theta_1 \leq (\theta - \theta_2)$ which minimizes the error sum of squares in stage 1. Therefore, in order to denote the fixed values of the previous stage as being different from the variable values in the cross product term let $z_{1i} = \theta_1 x_{1i}$. It should also be pointed out that $\sum_{i=1}^{n} y_i^2$ is simply a constant.

$$\therefore f_2(\theta) = \min_{0 \leq \theta_2 \leq \theta} \Big[\sum_{i=1}^{n} (\theta_2 x_{2i} - y_i)^2 - \sum_{i=1}^{n} y_i^2 + 2 \sum_{i=1}^{n} z_{1i}\, \theta_2 x_{2i} + f_1(\theta - \theta_2) \Big] .$$

It would be possible to write the general recursive relationship at this point, however, there is a subtle point that should be illustrated. This will be accomplished by considering the third stage or third term to be introduced into the model.

$$f_3(\theta) = \min_{0 \leq \theta_3 \leq \theta} \Big[\sum_{i=1}^{n} (\theta_1 x_{1i} + \theta_2 x_{2i} + \theta_3 x_{3i} - y_i)^2 \Big]$$

$$= \min_{0 \leq \theta_3 \leq \theta} \Big[\sum_{i=1}^{n} (\theta_1^2 x_{1i}^2 + 2\theta_1\, x_{1i}\, \theta_2\, x_{2i} + 2\theta_1\, x_{1i}\, \theta_3\, x_{3i} - 2\theta_1 x_{1i}\, y_i + \theta_2^2 x_{2i}^2 + 2\theta_2\, x_{2i}\, \theta_3\, x_{3i} - 2\theta_2 x_{2i}\, y_i + \theta_3^2 x_{3i}^2 - 2\theta_3 x_{3i}\, y_i + y_i^2) \Big]$$

However, $f_2(\theta) = \sum_{i=1}^{n} (\theta_1^2 x_{1i}^2 + \theta_2^2 x_{2i}^2 + y_i^2 + 2\theta_1 x_{1i}\theta_2 x_{2i}$

$$- 2\theta_1 x_{1i} y_i - 2\theta_2 x_{2i} y_i)$$

and $$\sum_{i=1}^{n} (\theta_3 x_{3i} - y_i)^2 = \sum_{i=1}^{n} (\theta_3^2 x_{3i}^2 - 2\theta_3 x_{3i} y_i + y_i^2)$$

$$\therefore\ f_3(\theta) = \min_{0 \leq \theta_3 \leq \theta} [\sum_{i=1}^{n} (\theta_3 x_{3i} - y_i)^2 - \sum_{i=1}^{n} y_i^2$$

$$+ 2 \sum_{i=1}^{n} \theta_3 x_{3i} (\theta_1 x_{1i} + \theta_2 x_{2i}) + f_2(\theta - \theta_3)]$$

The point that should be noted is that $(\theta_1 x_{1i} + \theta_2 x_{2i})$ is fixed for a given $(\theta - \theta_3)$ and this is the computed values of the model through the previous stage which minimized the error sum of squares for the specified sum of weights $(\theta_1 + \theta_2)$. Note, $(\theta - \theta_3) = (\theta_1 + \theta_2)$. Therefore, if these fixed values of $(\theta_1 x_{1i} + \theta_2 x_{2i})$ are represented as $z_{3i} = (\theta_1 x_{1i} + \theta_2 x_{2i})$ then;

$$f_3(\theta) = \min_{0 \leq \theta_3 \leq \theta} [\sum_{i=1}^{n} (\theta_3 x_{3i} - y_i)^2 - \sum_{i=1}^{n} y_i^2 + 2 \sum_{i=1}^{n} \theta_3 x_{3i} z_{2i}$$

$$+ f_2(\theta - \theta_3)]$$

This type of dynamic programming formulation requires that n additional values of z_{si} be carried from stage to stage. It does not require that they be retained for all previous stages only the preceding one since they are cumulative in nature. Then

$$z_{si} = \sum_{j=1}^{s} \theta_j x_{ji} , \quad i = 1, 2, \ldots, n.$$

for the computations of stage s. The n stage recursive relation can be written as:

$$f_n(\theta) = \min_{0 \leq \theta_n \leq \theta} \Big[\sum_{i=1}^{n} (\theta_n x_{ni} - y_i)^2 - \Sigma y_i^2 + 2 \sum_{i=1}^{n} \theta_n x_{ni} z_{ni}$$

$$+ f_{n-1}(\theta - \theta_n) \Big] , \quad n \geq 2$$

$$f_1(\theta) = \sum_{i=1}^{n} (\theta x_{1i} - y_i)^2$$

In order to compute these values by standard methods it will be necessary to make θ discrete. The increments can be refined to any level necessary in order to obtain a satisfactory weighting of the terms. This is essentially a one dimensional allocation problem with minor modifications.

Normally it would be expected that $\sum_{j=1}^{n} \theta_j = 1$ would be the constraint on θ; however, due to the built-in sensitivity of dynamic programming it may be desirable to constrain θ to a sum greater than 1 in order to determine if biases may be contained in the original individual terms. The dynamic programming solution will provide solutions for all values of θ. Further, it will also provide solutions for all arrangements of consecutive groupings of terms with the last term being deleted each time. By ordering the terms according to their suspected importance with respect to the model being constructed, the contribution of each added term can be considered for deletion in the reverse order in which it entered into the calculations. A simple example is constructed below for illustration of the method.

Consider the following example application of the dynamic programming technique discussed above. The data are as follow:

$$Y = [1, 3, 8]; X_1 = [1, 2, 3]; X_2 = [0, 1, 4]; X_3 = [0, 1, 3],$$

where Y is the dependent variable and the X_i are independent variables which are to be used as predictors of Y. It is further assumed that the functional form of each of the three predictors has been given as:

$$\hat{Y}_1 = \alpha_1 X_1 ; \hat{Y}_2 = \alpha_2 \sqrt{X_2} ; \hat{Y}_3 = \alpha_3 X_3^2 .$$

By using a least squares fit for these three equations the following values of α are obtained.

$$\hat{Y}_1 = 1.857 X_1 ; \hat{Y}_2 = 3.80 \sqrt{X_2} ; \hat{Y}_3 = 0.915 X_3^2 .$$

The problem is that of combining these three estimates into a single estimate which will reflect the nature of Y for extrapolated values of the independent variables. These data could be combined in a single equation by the least squares method which would simply be the solution of linear equalities which would provide

$$Y = 1X_1 + 5/7 \sqrt{X_2} + 2/7 X_3^2 .$$

However, if four terms were used, the least squares approach could not have been used. The solution of this problem by dynamic programming is given in Figure 1.

The solution provided by dynamic programming yields weights of 0.4, 0.4 and 0.2 for X_1, X_2 and X_3, respectively. This will provide a final model of the form,

$$\begin{aligned} \hat{Y} &= (0.4)\ 1.857 X_1 + (0.4)\ 3.80 \sqrt{X_2} + (0.2)\ 0.915 X_3^2 \\ &= 0.7428 X_1 + 1.52 \sqrt{X_2} + 0.1830 X_3^2 . \end{aligned}$$

θ	$f_{1(\theta)}$	z_{11}	z_{12}	z_{13}	θ_1	$f_{2(\theta)}$	z_{21}	z_{22}	z_{23}	θ_2	$f_{3(\theta)}$	z_{31}	z_{32}	z_{33}	θ_3
0	74.0000	0	0	0	0	74.0000	0	0	0	0	74.000	0	0	0	0
0.2	52.9043	.3714	.7428	1.1142	.2	48.0080	0	.7600	1.5200	.2	48.0080	0	.7600	1.5200	0
0.4	27.6484	.7428	1.4856	2.9712	.4	27.7920	0	1.5200	3.0400	.4	27.7920	0	1.5200	3.0400	0
0.6	13.1627	1.1142	2.2284	4.4568	.6	12.9469	.7428	2.2456	4.4912	.2	12.9469	.7428	2.2456	4.4912	0
0.8	4.4704	1.4856	2.9712	5.9424	.8	4.0215	.7428	3.0056	6.0112	.4	4.0215	.7428	3.0056	6.0112	0
1.0	1.5714	1.8571	3.7142	7.4284	1.0	0.8264	1.1142	3.7484	7.4968	.4	<u>0.2185</u>	.7428	3.1886	7.6582	.2
*1.2	4.4659	2.2284	4.4568	8.9136	1.2	3.2855	.3714	4.5428	8.7142	1.0	2.1595	.7428	2.7946	8.4322	.6

*Deviates from minimizing the error sum of squares in order to illustrate some advantages of the technique, normally the values for this row in the example would be the same as for $\theta = 1.0$.

Figure 1. Tabulation of Dynamic Programming Solution for Example Problem.

This would probably be considered to be different from the solution given by the least squares approach.

It probably should be reiterated that the dynamic programming procedure is primarily for the limited data case and is not offered as a replacement for standard statistical techniques.

PART II

QUANTIFICATION OF SUBJECTIVELY DETERMINED DATA IN THE FORMULATION AND UTILIZATION OF MATHEMATICAL MODELS

N67-16116

By

Glen Self

For

Industrial Engineering Department

Texas A&M University

November 1966

QUANTIFICATION OF SUBJECTIVELY DETERMINED DATA IN THE FORMULATION AND UTILIZATION OF MATHEMATICAL MODELS

by

Glen D. Self, Assistant Professor
Department of Industrial Engineering
Texas A&M University

Abstract

The limited data case in the construction of mathematical models has created an area of research for the utilization of subjectively determined data. Research at Texas A&M has pointed out some applications in cost modeling and has plans for extending the research. In general it has provided for a systematic methodology for model development in the extremely limited historical data case by quantification of expertise.

Introduction

Recently the use of subjectively determined data in planning and decision making functions has become more of an accepted and acknowledged approach in those areas where little or no historical data exists. Some of the more widely known developments are Industrial Dynamics at MIT, Project Delphi at RAND, PATTERN at Minneapolis - Honeywell and some recent work within NASA. One point of interest which has been established is that most everyone will provide an honestly considered opinion and will argue at lengths over whether a function is concave or convex and cite specific examples in defense of their respective positions.

Often the outcome of these discussions is the resolution of sematic differences.

In general the acceptance of the quantification of expertise has been positive in nature. There is a near infinity of published articles dealing with the quantitative problems that relate to the management decision making process. However, the area of developing the inputs necessary to use those tools is an often neglected area. The response that these subjective data methods have received is an indication of the wealth of information that experience provides as it relates to the formulation and use of mathematical models in decision making. The research presently being conducted at Texas A&M in this area is directed toward the further refinement of the subjective data approach. The research is concerned with utilizing a basically statistical approach to the problem of resolving questions of convergence, distinguishability or discrimination among subjectively assigned values, and the formulation of mathematical models. The details of the research at Texas A&M may be found in references (1) and (2). The specific work is oriented toward cost models; however, the approach and the research problems evolved are applicable to the general area of utilization of subjective data. This research is an outgrowth of some previous experiments in the area of quantification of expertise (3), (4).

The Model Builder's Dilemma

A number of organizations have initiated construction of mathematical models to represent various aspects of the space program in order that they might be used to assist in the planning function. As these models are expanded and become more complex in order to represent the detailed aspects of the process involved, the collection of historical data to support the formulation of the model soon forces the model builders to enter into an extensive campaign to collect data or to contract the dimensions of the model to those more compatible with the data available. In addition to the limited amount of detail available in terms of data, there is also the problem of the number of samples available with which a homogeneous data base can be formed. Launch vehicle programs are more numerous than the spacecraft programs and can provide one to five more data points on a subsystem basis. However, model building from the standpoint of using standard multiple regression techniques or other statistical inference techniques soon encounters difficulties when testing the significance of parameters derived from these limited data. Extrapolation of results may tend to cast doubt upon the validity of the model itself. The basis of the judgement of these models is the potential model user whose experience background indicates that future events would differ from those provided by the model. If in fact this information is accurate and can be quantified it could be of value in formulation or utilization of the model. The availability of persons with more detailed information that could be used to equate the various data collection systems and break out component parts of the total cost picture as the model builder has need, introduces the question of whether a large amount of good information is being passed over in order to obtain some questionable data.

In general the "no data" and "gross data" situation leads to various approaches in model development. This takes on different forms such as, (1) labeling the estimation or prediction function as one of managements prerogatives, (2) using gross models with the contention that the model will give satisfactory results, (3) installing data collection systems that will provide data for future model development, and (4) providing for the development of a logical and consistent model structure with the plan that all available data will be used with provisions for revision as more data becomes available and provisions for incorporating the available expertise. This latter course is the one primarily under investigation and will be the subject of this discussion.

The utilization of subjectively determined data has shown a number of uses and the results are still in the trial and acceptance stage, an academic look at these developments, results and projected results are felt to be in order so that an exchange of ideas may be promoted and this general area of research evaluated.

Development of Research Area

The basic ideas developed in this paper were initiated and initial development began with the Cost Implication Model (4) and the Contingency Planning Model (5). These models provided for the use of subjectively determined data in order to assess the effects of various program contingencies as deviations from a pre-established base case. In general, the models were to assist in the evaluation of time dependent occurrences which could not be accommodated in the normally static type cost model. This led to the construction of computerized models with library data that was highly flexible as to content and use within a pre-established model structure. The models provided for changes in both library data and problem data to determine the form of various cost-time relationships to be used in specific applications of the models.

The response of this type of model to the specific details of a planning situation and the generally acceptable results led to the further exploration and additional research in the area.

Many readers schooled in the classical modeling techniques will object on the basis of user bias being introduced directly into the model. The counter argument is that the models referred to above are simply the tools of the user and the results are his responsibility relative to the problem being explored. The purpose of the model is to combine in a logical and consistent manner all of the information available at the time the model is exercised. It is simply to assure that all things have been considered by the model user and to provide for the use of available subjective data along with existing historical data.

Generally some known data point is used and the general form of the extrapolation is left to expertise. In this way it is possible for the model user to contend that if the cost at some point in time is of a given level, then the future costs are expected to accumulate according to a specific pattern. Obviously, if the general form of the behavior of costs can be extrapolated into the future, then the historical data could be used to determine the relative location of a curve of that form. In this manner, limited data of three or more points can be used to an advantage without obtaining trivial answers. By trivial, it is meant that the model fits the data exactly and does not behave properly outside the range of the limited data.

The research in the general area of the limited data case has been along a number of different paths. However, most developments have been in the area of using subjective data to extend the results available from the more quantitative tools and then the use of quantitative tools within the subjective extensions of the available data. Research which has been completed in this area includes the topics discussed below.

Combining Estimators in the Limited Data Case

In the limited data case it is necessary to determine the "best fit" of a functional form to the available data by use of a single independent variable. If there are a number of variables which are generally considered to influence the response being studied then there is the additional problem of combining the individual predictors into a single predictor which can be considered "best" from the standpoint of the data available. Theoretically, if there is a single response which is predicted by a number of variables and each variable is represented independently by an equation, then each equation could be used separately to predict cost. Therefore, it is suggested that the individual predictors cannot be added but must be weighted to obtain a best estimate of the cost.

If they met the requirements of independent estimators with sufficient statistical information then the best linear combination could be formed by weighting them inversely proportional to their individual variances. However, the lack of data and independence suggests that some other method should be used. It is possible to apply a subjective weighting by selecting or ordering the predictor variables from the most important to the least important and assign weights which sum to unity. However, it was shown that a dynamic programming formulation of the problem could determine the best weights to be used in combining the functions, where best is to minimize the error sum of squares. This part of the research was used to point up the fact that once the modeling formulation has been turned over to the subjective mode, it is still possible to introduce quantitative tools into the formulation procedure.

The results of this particular formulation permits the model builder to take advantage of the built-in sensitivity of dynamic programming in selecting the values of the weighting coefficients and the number of terms to be included in the combined model.

Evaluation of Interaction through the Use of Expertise

The individual predictors in the preceding paragraph utilized expertise in the determination of the functional form of the individual cost predictors due to the lack of sufficient data to use multiple regression techniques. This area of research has extended the application of expertise into the evaluation of interaction effects and the incorporation of these terms into the model. This development is for the limited data case and permits the minimum sum of squares technique to be used as a basis for assessing the fit of the function. The specific research problem developed when the weighting of the individual predictors of cost were not reproducing the original data used in the formulation of the model. The reproduction of that data is not a difficult task since a polynomial of one degree less than the number of data points would fit exactly. However, the functional form of the main effects or first order effects was selected on a subjective data base. That is, the particular form of the costs, given that the single independent variable involved is increased over a wide range values, has been determined to be of some form, for example $y = a+bx$, $y = a \ln x$ or $y = ae^{bx}$. Therefore, when these are combined to obtain a single estimate of y, it is possible that the results will be unacceptable even for the historical levels of the variables.

These major deviations are assumed to be due to the interaction of the variables. If expertise is used to select the form of the second order terms, that is, those involving pairs of the independent variables, then it is possible to use the original data over again to determine a weighting of the second order terms in order to minimize the error sum of squares. The second order terms are formulated to represent the difference between the first order effects and the actual data. This is not to be confused with what would be the residual of a least squares fit, it is a number that has a mean not equal to zero and has the physical interpretation of interaction. This procedure can be repeated by the model builder as long as the functional form of the interaction has physical meaning. This area of research has demonstrated still another combination of quantitative tools and the quantification of subjective data to provide for a consistent and reproducible methodology for development of interaction terms for mathematical models which have a limited data base.

Since the behavior of a single equation of the type developed by the methods discussed above would be difficult to evaluate over a wide range of each of the variables, it was convenient to develop a computer program which would perform such evaluations. The results of this part of the study showed the total combined estimators and their interaction terms to be extremely well behaved over a wide range of values of the variables involved. The basic idea was to provide the model developer-user with a demonstration of its applicability over a wide range of values in order that it may be incorporated into a still larger model.

The total value of the research can be shown in the performance of a single equation of the form

$$y = a \ln x_2 + b \ln x_3 + c \ln x_4 + d \ln^2 x_1 + e \ln^2 x_2 + f \ln^2 x_3$$
$$+ g \ln^2 x_4 + h \ln x_1 \ln x_3 + p \ln x_1 \ln x_4 + q \ln x_2 \ln x_3$$
$$+ r \ln x_2 \ln x_4 + s \ln x_3 \ln x_4 + t \ln^3 x_1 + u \ln^3 x_2 + v \ln^3 x_3$$
$$+ w \ln^3 x_4 + x \ln x_1 \ln x_2 \ln x_3 \ln x_4$$

which was derived by using three data points and the expertise approach outlined above.

Estimation of Run-Out Costs for Spacecraft Programs

The area of estimation of run-out costs for on-going spacecraft programs is one where another merger of the quantitative tools and subjective data has been shown to be compatible. The development in this area was a computerized application of the best fit of the data from a partially completed subsystem to candidate curves of the percent cost-percent time type determined from historical data. The utilization of subjective data was for determining the algebraic form of the fitting function and establishing which curves taken from historical data should be considered as candidate curves for prediction of run-out costs of current subsystems. The quantitative tools are introduced in fitting a polynomial, computing a least squares criteria for fit of the curve to the candidate curves and projection of those costs according to the functional form of the curve selected. The computerized model for performing the total analysis has been completed.

Additional research is being considered in the area of adaptive type models which can automatically incorporate new data into the system and change form as required in consideration of previously determined data. Investigations into self-adapting model structure and library data concepts are being considered. This will necessarily include the mixture of expertise, historical data, and data with a subjective estimate of its accuracy in perspective to the formulation, growth and utilization of mathematical models.

Collection and Analysis of Subjectively Determined Data

An extensive discussion could be presented relative to the collection and analysis of expertise; however, only the approach presently being used for collection of data relating to percent time versus percent cost of spacecraft subsystems will be discussed.

The basic plan in the development of this area has been one of requesting experts to express their opinion as to the behavior of percent cost-percent time upon the basis of their experience, primarily with the Gemini program. The first questionnaire provided complete freedom in specification of the form of the cost-time relationship. These data were analyzed to determine a set of general curves which were submitted with the second questionnaire. In this case the experts selected a single curve from those provided to describe the cost-time relationship for each of one hundred cost categories. The selections are then tested to determine if the selections were non-random on a per category basis by use of multinomial computations. If this provides a non-random indication, then a composite of the curves selected for each category is determined by a weighted average. These data provide an estimate of the cost variance over the range of time by direct computation.

If the selections are shown to be random in nature, then more information is supplied the experts and the tests repeated. If agreement cannot be achieved then deletion of subjective data by use of extreme value techniques is initiated.

This research is directed toward determining the sensitivity of the model which uses these data. That is, the estimates of variance will be used to indicate the possible error that could be obtained using the subjectively determined data. These estimates will be used to evaluate the sensitivity of the model to errors in the expertise. However, it should be pointed out that this is not unlike the possible error that could be introduced by the use of historical data for building the model.

This particular concept has possible applications in a number of other areas and will be extended as the opportunities present themselves.

Summary and Conclusions

Quantification of subjective data has been associated with the field of operations research for a number of years in the area of utility theory. The experimentation in the determination of utiles or value associated with various programs has met with varied results. Perhaps this failure or success of these investigations has been dependent upon the quantitative tools or methods which have accompanied the subjectiveness. The rigor of statistics, mathematics, various optimization techniques and significance tests have something to contribute to the field of quantification of expertise. The technique itself is seen as a logical development in the extension of complex models to more closely approximate the total information available to the decision maker. The intent of this research is not directed at making the decision for the manager or usurping his prerogatives. It is directed at the systematic consideration of all aspects of the problem one factor at a time and the structuring of the analysis such that each of the elements will be combined in the appropriate manner to provide the integrated analysis desired.

Generally, management eventually makes some type of decision that results in the issuance of that decision in terms of quantitative values. It was the objective of the research reported and future research planned in this area to introduce the best quantitative techniques applicable into the problem of quantifying subjective data and subsequently using that data in the formulation and utilization of mathematical models. Often these models will be computerized and will make the utilization of that data more convenient. Admittedly these techniques are in the development stage, but one benefit that may be derived from research of this type is that it may cause the collection of actual data of the type needed for the formulation of mathematical models to be expedited. If this is accomplished then the utilization of subjective data will have achieved a major contribution in the field of mathematical models.

It should be pointed out that the research reported in this article was using expertise to represent physical functions, not to represent emotion. If it is physical functions that are being represented by expertise, then they can be tested with the accumulation of actual data and will serve in the interim as the best that is available for the mathematical model. A direct comparison of the type of results obtained in the discussion presented above and the type of data presently available for formulation of the models being considered might provide a more realistic attitude toward the subject area and adjust the evaluation of its worth at this time.

One area that has been suggested for the use of the models that use the subjective data base is in the area of management training, or the training of project leaders. The experience of previous project managers could be placed in the form of a model that represents their experience in various physical situations in the past, this could be used as a project leader "game" for training.

In summary it can be said that the research reported in references (1) and (2) has precipitated thought on the subject as well as provided some insight into the problems of cost estimation under the limited data case.

References

(1) "The Determination of Empirical and Analytical Spacecraft Parametric Curves" Progress Report I, NASA Grant SC-NGR-44-001-027, Texas A&M Research Foundation, December 1, 1965.

(2) "The Determination of Empirical and Analytical Spacecraft Parametric Curves" Progress Report II, NASA Grant SC-NGR-44-001-027, Texas A&M Research Foundation, May 1, 1966.

(3) "A Study of Mission Requirements for Manned Mars and Venus Exploration," General Dynamics/Fort Worth, FZM-4366-3, May 30, 1965.

(4) "Launch Vehicle Cost Parameters Study," General Dynamics/Fort Worth, FZM 4247-4, May 27, 1965.

(5) "Contingency Planning Model," General Dynamics/Fort Worth, FZM 4458-1, August 30, 1965.

N67-16117

PART III

WEIGHTED ESTIMATION OF RUN-OUT COSTS

By

Sidney P. Brown

For

Industrial Engineering Department

Texas A&M University

November 1966

INTRODUCTION

Various methods have been used to estimate the total cost of space programs at various points during the life of the program. The previous report contains one of these ways, but for various reasons (including a request from NASA MSC), a continuation to extend the research into this area was made. The problem of selecting a cost distribution function over time from among many as a best fit to a partial cost distribution, and the use of the one selected to extrapolate the partial costs to a completed cost was originated for the utilization of Gemini data for the extrapolation of Apollo data. However, the generalization of the problem is apparent, and in fact has already been used for cost estimates of a moon t.v. camera.

Since this is an extension of previously reported research, there is necessarily some repetition of material for orientation purposes, but the addition merits the repetition.

The hypothesis was that a subsystem of a completed project would follow the percent time vs. percent cost curve of some subsystem of a previously completed project. The subsystems of the uncompleted project would not necessarily follow the curve of the same subsystem of the completed project because of such factors as the amount of parallel development, technological difficulties, program changes and other similar reasons. Therefore, the choice of the best fitting curve should be made from among a "population" of curves of subsystems of completed projects. Then the curve which approximates the available data the closest should be used to estimate the run-out cost.

There are three basic phases to the estimation of run-out costs in this method:

(1) The determination of a polynomial to fit each "population curve.

(2) The determination of the best fitting curve among the population to the uncompleted subsystem data and

(3) The determination of the run-out costs. These phases are shown on the flow chart of Figure 1.

DETERMINE POLYNOMIAL

The data of a completed project which was available to Texas A&M was the percent time vs. percent cost curves of Gemini. This data consisted of four intermediate points through which a smooth curve had been drawn (NASA/MSC data).

To obtain an equation of the different curves, a third degree polonomial was determined to be the best general fit to the curves. The general form of the equation is:

$$f(x) = AX + BX^2 + CX^3$$

To determine the coefficients of this equation, the method of least squares was used. A set of four simultaneous normal equations, involving summations of the decimal percent time raised to powers from zero to six and the same summations except being multiplied by percent cost with the powers ranging from zero to three, is solved to give the desired coefficients. The more data points that are used in the computations, the better the fit is expected to be. The degree of the polynomial could be increased if desired in order to obtain a still better fit. This technique was used to determine the equation and was found to give satisfactory results.

DETERMINE POLYNOMIAL

FOR EACH OF THE SIGNIFICANTLY DIFFERENT COMPLETED SUBSYSTEM CURVES, DETERMINE A THIRD DEGREE POLYNOMIAL THAT BEST FITS THE AVAILABLE DATA POINTS

DETERMINE BEST FITTING CURVE

FOR EACH UNCOMPLETED SUBSYSTEM, DETERMINE THE BEST FIT OF THE AVAILABLE DATA POINTS TO EACH COMPLETED SUBSYSTEM CURVE AND CHOOSE THE ONE THAT BEST FITS

DETERMINE RUN-OUT COSTS

USING THE BEST FITTING CURVE, ANY COST DATA POINT CAN BE USED TO DETERMINE THE RUN-OUT COST.

FIGURE 1. FLOW CHART OF RUN-OUT COST ESTIMATION PROCESS

For the specific problem of runout costs of Apollo, the problem of distinguishing which of the cost distributions are significantly different. This is not a particularly difficult problem if the number of samples and conditions of independence in these samples are met; however, due to the nature of the data this is not a straight-forward application of statistical inference type of test of significance.

DETERMINE BEST FITTING GEMINI CURVE FOR APOLLO DATA

The data available for an uncompleted subsystem is given as a cost at a certain time (NASA/MSC data). From the data, the percent time of the project is known since the project length is known. What is needed is the percent cost each of the points represent.

The data points must be tested against each different completed system curve in such a way as to get the best fit and then choose the one curve which gives the best fit.

Since the ratios of the cost data points are known, it is possible to place the first intermediate point at a percent cost of X. With n intermediate data points of an uncompleted subsystems, the n-1 remaining data points are at a height k_2X, $k_3X, \ldots, k_nX$ (see Figure 2).

Using the method of least squares to provide the best fit,

$$(1) \quad (Y_1-X)^2 + (Y_2-K_2X)^2 + \ldots + (Y_n-k_nX)^2 = \text{minimum}$$

Where Y_1, Y_2, ..., Y_n are the values of the percent cost of the completed subsystem at the percent time of the data points of the uncompleted subsystem, these values are obtained from the derived polynomials of the completed subsystem.

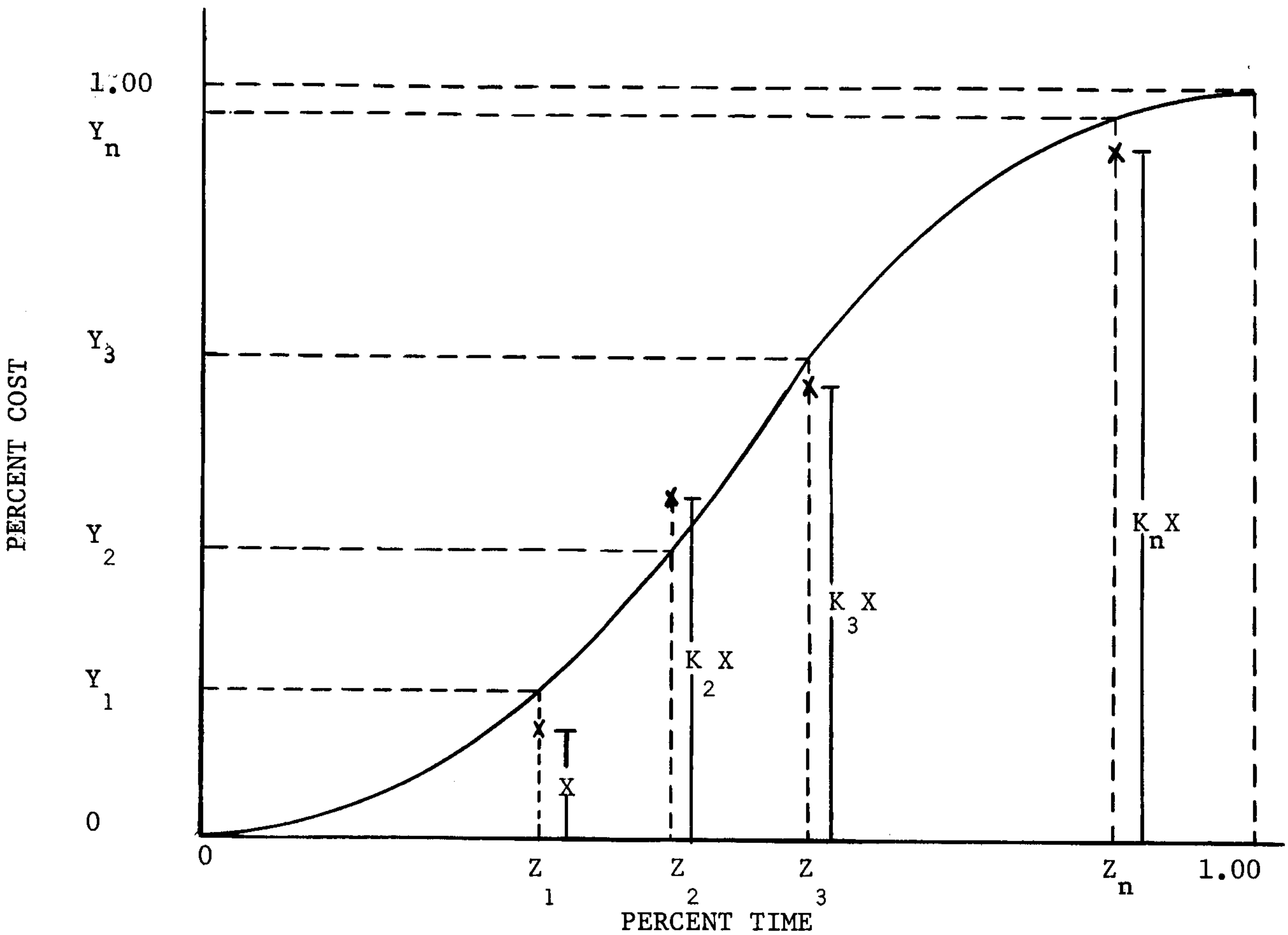

Figure 2. METHOD OF COST RATIO DETERMINATION

However, the usual method of least squares necessarily weighs each deviation the same regardless of where it occurs along the abcissa. There are reasonable arguments why the initial data points might not be as accurate as subsequent data points for reasons such as the accounting system not being well established.

More than that, it is intuitively appealing that as the project nears completion, these data points should be more indicative of the final runout cost than earlier ones.

A weighing system which is both appealing and easily used would be to weigh the error squared by the fraction of project time for each data point (Z_n').

The modified method of least squares to provide the best fit is given by (2).

$$(2)\quad [Z_1(Y_1 - X)]^2 + [Z_2(Y_2 - K_2X)]^2 + \ldots + [Z_n(Y_n - K_nX)]^2 = \text{min.}$$

Taking the derivative of (2) and setting it equal to zero to obtain the minimum, (2) $2(y_1 - X)Z_1\ (-Z_1) + 2(Y_2 - k_2X)Z_2\ (-k_2Z_2) + 2(Y_3 - k_3X)$

$$Z_3\ (-k_3Z_3) + \ldots + 2(Y_n - k_nX)Z_n(-k_nZ_n) = 0$$

Solving for X;

$$(3)\quad X = \frac{Y_1Z_1^2 + k_2Y_2Z_2^2 + k_3Y_3Z_3^2 + \ldots + k_nY_nZ_n^2}{Z_1^2 + k_2^2Z_2^2 + k_3^2Z_3^2 + \ldots + k_n^2Z_n^2}$$

By knowing the value of X, the value of the sum of the squares of (1) can be determined to obtain a measure of fit for the uncompleted subsystem to one of the completed subsystem curves. Repeating this with each of the completed subsystem curves gives a weighted measure of fit for each of these curves.

By using this measure of fit, the best fitting curve can be determined by choosing the curve which had a minimum value for the weighted sum of the squares.

DETERMINE RUN-OUT COST

With the best fitting completed subsystem curve chosen, and the uncompleted subsystem data points located, the run-out cost can be estimated.

The method used to choose the best curve was to place the data points of the uncompleted subsystem in the appropriate perspective to the completed subsystem. Each of these points were converted to a percent cost; therefore, each data point is a percentage of the run-out cost, and any one may be used to obtain a projection of the 100% or run-out cost of that subsystem.

Since any point may be used to estimate the run-out cost, the first point will be chosen for convenience since equation (3) located the first intermediate cost point at a percentage of the run-out cost, the relationship of the cost associated with that point to the projected cost is known. Thus, the run-out cost is the cost at the first point divided by the decimal percent of run-out cost (X).

A program which does the complete analysis of cost run-out has been completed. The data of the completed significantly different subsystem and the uncompleted subsystem is the only required information, with the estimated run-out cost as the information provided the user.

For various reasons, a cost projection for some point other than completed project time may be required. Therefore, the requested capability has been incorporated into the program. Program decks have been made available to M.S.C.

CARD 1

INPUT FORMAT COLUMN IDENTIFICATION

NOTE: A 1 indicates an option is to be used, a 0 or blank causes the program to delete an output option. All fields are Format I5 (right adjusted)

K9 - Option 9, Number of intermediate cost points desired

K8 - Option 8, Number of copies of output desired

K7 - Option 7, Output of curve name which data best fits Available only with Option 1

K6 - Option 6, Output of sum of squares information

K5 - Option 5, Output of uncompleted subsystem name

K4 - Option 4, Output of comparison between actual and computed points

K3 - Option 3, Output of cost equation

K2 - Option 2, Output of simultaneous equations that are solved for cost equation coefficients

K1 - Option 1, Output of completed subsystem name

M - Number of data points for each uncompleted subsystem

L - Number of data points for each completed subsystem

K - Number of completed subsystems to be used

PROBLEM DATA CARD 1

CARDS 2(a), 3(a), 4(a), ... , KC + 1(a) (Where C is integer $\leq$[L/4])

INPUT FORMAT COLUMN IDENTIFICATION

NOTE: TO BE INCLUDED ONLY IF OPTION ONE IS USED.

CARD IS TO BE INCLUDED IF COLUMN 20 ON FIRST DATA CARD HAS A 1 AND CARD IS TO BE OMITTED IF COLUMN 20 OF FIRST DATA CARD IS A ZERO OR BLANK

NAME OF COMPLETED SUBSYSTEM

DATA CARD SET 2

CARDS 2(b), 3(b), 4(b), ... , KC + 1(b)

INPUT FORMAT COLUMN IDENTIFICATION

Y(I + 3) - $(I + 3)^{th}$ decimal percent of cost of completed subsystem. Floating point.
X(I + 3) - $(I + 3)^{th}$ decimal percent of program time of completed subsystem. Floating point.
Y(I + 2) - $(I + 2)^{th}$ decimal percent of cost of completed subsystem. Floating point.
X(I + 2) - $(I + 2)^{th}$ decimal percent of program time of completed subsystem. Floating point.
Y(I + 1) - $(I + 1)^{th}$ decimal percent of cost of completed subsystem. Floating point.
X(I + 1) - $(I + 1)^{th}$ decimal percent of program time of completed subsystem. Floating point.
Y(I) - I^{th} decimal percent of cost of completed subsystem. Floating point.
X(I) - I^{th} decimal percent of program time of completed subsystem. Floating point.

DATA CARD SET 3

(AS MANY CARDS USED AS NEEDED TO CONTAIN EACH DATA POINT OF THE COMPLETED SUBSYSTEM)

CARDS KC + 2(a), KC + 3(a), KC + 4(a), ...

INPUT FORMAT COLUMN IDENTIFICATION

NOTE: To be included only if option five is taken.
Card is to be included if column 40 of first
data card is 1 and card is to be omitted if column
40 of first data card is 0 or blank.

NAME OF UNCOMPLETED SUBSYSTEM

DATA CARD SET 4

CARDS KC + 2(b), KC + 3(b), KC + 4(b), ...

INPUT FORMAT COLUMN IDENTIFICATION

C(I + 3) - $(I + 3)^{th}$ cost of uncompleted subsystem. Floating Point.
Z(I + 3) - $(I + 3)^{th}$ decimal percent of program time of uncompleted subsystem. Floating point.
C(I + 2) - $(I + 2)^{th}$ cost of uncompleted subsystem. Floating point.
Z(I + 2) - $(I + 2)^{th}$ decimal percent of program time of uncompleted subsystem. Floating point.
C(I + 1) - $(I + 1)^{th}$ cost of uncompleted subsystem. Floating point.
Z(I + 1) - $(I + 1)^{th}$ decimal percent of program time of uncompleted subsystem. Floating point.
C(I) - I^{th} cost of uncompleted subsystem. Floating point.
Z(I) - I^{th} decimal percent of program time of uncompleted subsystem. Floating point.

DATA CARD SET 5

(As many cards used as needed to contain each data point of the uncompleted subsystem)

CARDS KC + 2(c), KC + 3(c), KC + 4(c), ...

Note: To be included only if K9 is greater than zero. All fields are format F10.0.

INPUT FORMAT COLUMN IDENTIFICATION

XP(I+7) - (I + 7)th decimal percent of program time where estimate of cost of incompleted subsystem is desired.
XP(I + 6) - (I + 6)th "
XP(I + 5) - (I + 5)th "
XP(I + 4) - (I + 4)th "
XPC(I + 3) - (I + 3)th "
XP(I + 2) - (I + 2)th "
XP(I + 1) - (I + 1)th dec "
XP(I) - Ith decimal percent of program time where estimate of cost of uncompleted subsystem is desired.

DATA CARD SET 6

(As many cards used as needed to contain each point of the desired intermediate cost estimates)

PART IV

QUANTIFICATION AND UTILIZATION OF SUBJECTIVELY DETERMINED DATA IN THE CONSTRUCTION OF MATHEMATICAL MODELS

By

Grady Lee Haynes

(Under separate cover)

For

Industrial Engineering Department

Texas A&M University

November 1966

MAIN PROGRAM

```
      DIMENSION P(4,5), X(101), Y(101), Z(101),D(101),  C(101), YA(101),
     1YRO(101),IPIV(4),R1(100),R2(100),R3(100),RATIO(101),SLS(101),
     2W(100,101),NAME(101,5),NAMA(5),IREC(14),XP(50)
      REWIND 2
    4 IF (EOF(5)) 2,3,2
    3 READ (5,1) IREC
      WRITE (2,1) IREC
      GO TO 4
    1 FORMAT (13A6,A2)
    2 ENDFILE 2
      REWIND 2
      ICOPY = 1
   66 READ (2,13) K,L,M,K1,K2,K3,K4,K5,K6,K7,K8,K9
   13 FORMAT (12I5)
      WRITE (6,33)
   33 FORMAT (1H1,56X,20HCOMPLETED SUBSYSTEMS,//)
      DO 30 J=1,K
      IF (J-1)141,141,142
  142 WRITE (6,143)
  143 FORMAT (1H1)
  141 IF (1-K1) 14,15,14
   15 READ (2,16) (NAME (J,I),I= 1,5)
   16 FORMAT (5A6)
      WRITE (6,17) J, (NAME (J,I),I=1,5)
   17 FORMAT (//,64X,1H(,I3,1H),/,51X,5A6)
   14 P(1,1) = L
      P(1,2)=0
      P(1,3)= 0
      P(1,4) =0
      P(2,4)= 0
      P(3,4)= 0
      P(4,4)= 0
      P(1,5)= 0
      P(2,5)= 0
      P(3,5)= 0
      P(4,5)= 0
      READ (2,12)(X(I), Y(I),I=1,L)
   12 FORMAT (8F10.0)
      DO 50 I=1,L
      P(1,2) = P(1,2) +X(I)
      P(1,3)= P(1,3) + X(I)**2
      P(1,4) = P(1,4) + X(I)**3
      P(2,1) = P(1,2)
      P(2,2) = P(1,3)
      P(2,3) = P(1,4)
      P(2,4) = P(2,4) + X(I)**4
      P(3,1) = P(2,2)
      P(3,2) = P(2,3)
      P(3,3) = P(2,4)
      P(3,4) = P(3,4) + X(I)**5
      P(4,1) = P(3,2)
      P(4,2) = P(3,3)
      P(4,3) = P(3,4)
      P(4,4) = P(4,4) + X(I)**6
      P(1,5) = P(1,5) + Y(I)
      P(2,5) = P(2,5) + Y(I)*X(I)
      P(3,5) = P(3,5) + Y(I)*X(I)**2
```

```
50 P(4,5) = P(4,5) + Y(I)*X(I)**2
   IF (1-K2) 81,80,81
80 WRITE (6,36)
36 FORMAT (//,23X,86HTHE SET OF SIMULTANEOUS EQUATIONS THAT GIVES THE
  1 COEFFICIENTS OF THE COST EQUATION IS,)
   DO 38 I=1,4
   WRITE (6,37) P(I,1),P(I,2),P(I,3), P(I,4), P(I,5)
37 FORMAT (//,5X,E15.8,5H A + ,E15.8,5H B + ,E15.8,5H C + ,E15.8,
  25H D = , E15.8)
38 CONTINUE
81 DET = SIMEQN(P,IPIV,4,4,1.0E-15)
   R1(J) = P(2,5)
   R2(J) = P(3,5)
   R3(J) = P(4,5)
   IF (1-K3) 83,82,83
82 WRITE (6,51) P(2,5),P(3,5), P(4,5)
51 FORMAT (//,49X,34HTHE EQUATION OF THE COST CURVE IS,,/,27X,
  3 6HYA = (,E15.8,8H)(X) + (,E15.8,11H)(X**2) + (,E15.8,7H)(X**3))
83 IF (1-K4) 30,84,30
84 DO 75 I=1,L
   YA(I) = P(2,5)*X(I) + P(3,5)*X(I)**2 + P(4,5)*X(I)**3
   WRITE (6,55) I, YA(I), I, Y(I)
55 FORMAT (//,57X,3HYA(,I3,3H) =, F8.4,/,57X,3HY (,I3,3H) =,F8.4)
75 CONTINUE
30 CONTINUE
   WRITE (6,26)
26 FORMAT (1H1,55X,22HUNCOMPLETED SUBSYSTEMS,//)
59 IF(1-K5)58,111,58
111 READ (2,108) NAMA
108 FORMAT (5A6)
   WRITE (6,96) NAMA
96 FORMAT (//,51X,5A6)
58 READ (2,18)(Z(I), C(I),I=1,M)
18 FORMAT (8F10.0)
   DO 31 J=1,K
   DO 25 I=1,M
   W(J,I) = R1(J)*Z(I) + R2(J)*Z(I)**2 + R3(J)*Z(I)**3
25 CONTINUE
   RA=0
   RAT = 0
   DO 64 I=1,M
   D(I) = C(I)/C(1)
   RA = RA + D(I)*W(J,I)*Z(I)**2
   RAT = RAT + D(I)**2*Z(I)**2
64 CONTINUE
   RATIO(J) = RA/RAT
   SLS(J) = 0
91 DO 61 I= 1,M
   SLS(J) = SLS(J) +((W(J,I) -D(I)*RATIO(J))*Z(I))**2
61 CONTINUE
31 CONTINUE
   IF (1-K6) 34,90,34
90 WRITE (6,45)
45 FORMAT (//,34X,62HTHE SUM OF LEAST SQUARES USING THE I-TH COMPLETE
  1D SUBSYSTEM IS,)
   DO 32 J=1,K
92 WRITE (6,44) J,SLS(J)
44 FORMAT (//,53X,4HSLS(,I3,3H) =,E15.8)
32 CONTINUE
34 XHI = SLS(1)
   IB = 1
```

```
      DO 101 I=2,K
      IF (XHI.LE.SLS(I)) GO TO 101
      XHI = SLS(I)
      IB = I
  101 CONTINUE
      IF (1-K7) 93,22,93
   22 WRITE (6,102) IB,(NAME(IB,I),I=1,5)
  102 FORMAT (//,44X, 39HTHE CURVE WHICH BEST FITS THE DATA IS (,I3,2H),
     2/, 51X, 5A6)
      IF (K8) 150,93,150
  150 READ (2,153) (XP(I), I=1,K8)
  153 FORMAT (8F10.0)
      DO 152 I=1,K8
      WA = B1(IB)*XP(I) + B2(IB)*XP(I)**2 + B3(IB)*XP(I)**3
      ROCI = C(1)/RATIO(IB)*WA
      WRITE (6,155) I, ROCI
  155 FORMAT (//,38X,8HTHE NO. ,I1,23H INTERMEDIATE COST IS $,F12.1)
  152 CONTINUE
   93 ROC = C(1 )/RATIO(IB)
      WRITE (6,103) ROC
  103 FORMAT (//,49X,21HTHE RUN-OUT COST IS $,F12.1)
      IF(EOF(2)) 120,58,120
  120 REWIND 2
      IF(ICOPY.GE.K8 + 1) GO TO 130
      ICOPY = ICOPY +1
      GO TO 66
  130 STOP
      END
DATA
```

SAMPLE DATA LISTING

```
ATA
  7      6      4      1      1      1      1      1      1      1     30      2
    STRUCTURE (PRIMARY)
235      6096.0      0.43        10157.0     0.53        11852.0     0.625       13082.0
    0.90        0.95
      INSTRUMENTATION
235      1970.0      0.43        3665.0      0.53        4456.0      0.625       5024.0
    0.90        0.95
      COMMUNICATIONS
235      3934.0      0.43        5899.0      0.53        6685.0      0.625       7361.0
    0.90        0.95
     REACTION CONTROL
235      5870.0      0.43        11149.0     0.53        14134.0     0.625       17611.0
    0.90        0.95
       LAUNCH ESCAPE
235      936.0       0.43        1902.0      0.53        2452.0      0.625       2888.0
    0.90        0.95
    STABILITY CONTROL
0        0.0         0.285       0.37        0.48        0.61        0.58        0.72
675      0.81        1.0         1.0
  GUIDANCE AND NAVIGATION
0        0.0         0.285       0.387       0.48        0.609       0.58        0.708
675      0.778       1.0         1.0
   LANDING AND RECOVERY
0        0.0         0.285       0.20        0.48        0.41        0.58        0.546
675      0.718       1.0         1.0
     ELECTRICAL POWER
0        0.0         0.285       0.269       0.48        0.582       0.58        0.74
675      0.85        1.0         1.0
   ENVIRONMENTAL CONTROL
0        0.0         0.242       0.29        0.53        0.60        0.628       0.723
67       0.775       1.0         1.0
          PROPULSION
0        0.0         0.285       0.27        0.48        0.61        0.58        0.745
675      0.828       1.0         1.0
     STRUCTURE (ADAPTER)
0        0.0         0.285       0.396       0.48        0.698       0.58        0.788
675      0.863       1.0         1.0
```

(1)
STABILITY CONTROL

THE SET OF SIMULTANEOUS EQUATIONS THAT GIVES THE COEFFICIENTS OF THE COST EQUATION IS,

0.60000000E 01 A + 0.30200000E 01 B + 0.21036500E 01 C + 0.16363999E 01 D = 0.35100000E 01

0.30200000E 01 A + 0.21036500E 01 B + 0.16363999E 01 C + 0.13804407E 01 D = 0.23626000E 01

0.21036500E 01 A + 0.16363999E 01 B + 0.13804407E 01 C + 0.12331224E 01 D = 0.17818615E 01

0.16363999E 01 A + 0.13804407E 01 B + 0.12331224E 01 C + 0.11454203E 01 D = 0.14656199E 01

THE EQUATION OF THE COST CURVE IS,
YA = (0.12734183E 01)(X) + (0.24018385E 00)(X**2) + (-0.51360290E 00)(X**3)

YA(1) = 0.
Y (1) = 0.

YA(2) = 0.3705
Y (2) = 0.3700

YA(3) = 0.6098
Y (3) = 0.6100

YA(4) = 0.7192
Y (4) = 0.7200

YA(5) = 0.8110
Y (5) = 0.8100

YA(6) = 1.0000
Y (6) = 1.0000

THE SET OF SIMULTANEOUS EQUAT

0.60000000E 01 A +	0.30200000E 01 B +	0.2103
0.30200000E 01 A +	0.21036500E 01 B +	0.1636
0.21036500E 01 A +	0.16363999E 01 B +	0.1380
0.16363999E 01 A +	0.13804407E 01 B +	0.1233

THE

YA = (0.15325031E 01)(X)

(2)
AND NAVIGATION

GIVES THE COEFFICIENTS OF THE COST EQUATION IS,

+ 0.16363999E 01 D = 0.34820000E 01

+ 0.13804407E 01 D = 0.23384050E 01

+ 0.12331224E 01 D = 0.17643951E 01

+ 0.11454203E 01 D = 0.14537200E 01

OF THE COST CURVE IS,
)45120E 00)(X**2) + (0.47957535E-01)(X**3)

1) = 0.
1) = 0.

2) = 0.3907
2) = 0.3870

3) = 0.6072
3) = 0.6090

4) = 0.7029
4) = 0.7080

5) = 0.7847
5) = 0.7780

6) = 1.0000
6) = 1.0000

LANDING AND RECOVERY

THE SET OF SIMULTANEOUS EQUATIONS THAT GIVES THE COEFFICIENTS OF THE COST EQUATION IS,

```
0.60000000E 01 A +  0.30200000E 01 B +  0.21036500E 01 C +  0.16363999E 01 D =  0.28740000E 01

0.30200000E 01 A +  0.21036500E 01 B +  0.16363999E 01 C +  0.13804407E 01 D =  0.20551300E 01

0.21036500E 01 A +  0.16363999E 01 B +  0.13804407E 01 C +  0.12331224E 01 D =  0.16215221E 01

0.16363999E 01 A +  0.13804407E 01 B +  0.12331224E 01 C +  0.11454203E 01 D =  0.13773223E 01
```

THE EQUATION OF THE COST CURVE IS,

```
YA = ( 0.67917220E-01)(X) + ( 0.24107553E 01)(X**2) + (-0.14789307E 01)(X**3)
```

```
YA(  1) =  0.
Y (  1) =  0.

YA(  2) =  0.1809
Y (  2) =  0.2000

YA(  3) =  0.4245
Y (  3) =  0.4100

YA(  4) =  0.5618
Y (  4) =  0.5460

YA(  5) =  0.6894
Y (  5) =  0.7180

YA(  6) =  0.9997
Y (  6) =  1.0000
```

THE SET OF SIMULTANEOUS EQUAT

0.60000000E 01 A + 0.30200000E 01 B + 0.2103

0.30200000E 01 A + 0.21036500E 01 B + 0.1636

0.21036500E 01 A + 0.16363999E 01 B + 0.1380

0.16363999E 01 A + 0.13804407E 01 B + 0.1233

THE

YA = (0.29594864E 00)(X)

(4)
RICAL POWER

GIVES THE COEFFICIENTS OF THE COST EQUATION IS,

+ 0.16363999E 01 D = 0.34410000E 01

+ 0.13804407E 01 D = 0.23589750E 01

+ 0.12331224E 01 D = 0.17921596E 01

+ 0.11454203E 01 D = 0.14763893E 01

OF THE COST CURVE IS,
74474E 01)(X**2) + (-0.22933761E 01)(X**3)

1) = 0.
1) = 0.

2) = 0.2747
2) = 0.2690

3) = 0.5790
3) = 0.5820

4) = 0.7325
4) = 0.7400

5) = 0.8602
5) = 0.8500

6) = 1.0000
6) = 1.0000

ENVIRONMENTAL CONTROL

THE SET OF SIMULTANEOUS EQUATIONS THAT GIVES THE COEFFICIENTS OF THE COST EQUATION IS,

```
0.60000000E 01 A +  0.30700000E 01 B +  0.21827480E 01 C +  0.17114856E 01 D =  0.33880000E 01

0.30700000E 01 A +  0.21827480E 01 B +  0.17114856E 01 C +  0.14393845E 01 D =  0.23614740E 01

0.21827480E 01 A +  0.17114856E 01 B +  0.14393845E 01 C +  0.12753403E 01 D =  0.18185607E 01

0.17114856E 01 A +  0.14393845E 01 B +  0.12753403E 01 C +  0.11741656E 01 D =  0.15055952E 01
```

THE EQUATION OF THE COST CURVE IS,

```
YA = ( 0.10931744E 01)(X) + ( 0.38231005E 00)(X**2) + (-0.47605430E 00)(X**3)
```

```
YA(  1) =  0.
Y (  1) =  0.

YA(  2) =  0.2802
Y (  2) =  0.2900

YA(  3) =  0.6159
Y (  3) =  0.6000

YA(  4) =  0.7194
Y (  4) =  0.7230

YA(  5) =  0.7609
Y (  5) =  0.7750

YA(  6) =  0.9994
Y (  6) =  1.0000
```

(6)
PROPULSION

THE SET OF SIMULTANEOUS EQUATIONS THAT GIVES THE COEFFICIENTS OF THE COST EQUATION IS,

```
0.60000000E 01 A +  0.30200000E 01 B +  0.21036500E 01 C +  0.16363999E 01 D =  0.34530000E 01

0.30200000E 01 A +  0.21036500E 01 B +  0.16363999E 01 C +  0.13804407E 01 D =  0.23607499E 01

0.21036500E 01 A +  0.16363999E 01 B +  0.13804407E 01 C +  0.12331224E 01 D =  0.17903502E 01

0.16363999E 01 A +  0.13804407E 01 B +  0.12331224E 01 C +  0.11454203E 01 D =  0.14737186E 01
```

THE EQUATION OF THE COST CURVE IS,

```
YA = ( 0.45990701E 00)(X) + ( 0.25680172E 01)(X**2) + (-0.20274310E 01)(X**3)
```

```
YA(  1) =  0.
Y (  1) =  0.

YA(  2) =  0.2927
Y (  2) =  0.2700

YA(  3) =  0.5882
Y (  3) =  0.6100

YA(  4) =  0.7351
Y (  4) =  0.7450

YA(  5) =  0.8570
Y (  5) =  0.8280

YA(  6) =  1.0005
Y (  6) =  1.0000
```

STRUCTURE (ADAPTER)

THE SET OF SIMULTANEOUS EQUATIONS THAT GIVES THE COEFFICIENTS OF THE COST EQUATION IS,

```
0.60000000E 01 A +  0.30200000E 01 B +  0.21036500E 01 C +  0.16363999E 01 D =  0.37449999E 01

0.30200000E 01 A +  0.21036500E 01 B +  0.16363999E 01 C +  0.13804407E 01 D =  0.24874649E 01

0.21036500E 01 A +  0.16363999E 01 B +  0.13804407E 01 C +  0.12331224E 01 D =  0.18512719E 01

0.16363999E 01 A +  0.13804407E 01 B +  0.12331224E 01 C +  0.11454203E 01 D =  0.15055215E 01
```

THE EQUATION OF THE COST CURVE IS,

```
YA = ( 0.14123813E 01)(X) + ( 0.36172765E 00)(X**2) + (-0.77351250E 00)(X**3)
```

```
YA(  1) =  0.
Y (  1) =  0.

YA(  2) =  0.4140
Y (  2) =  0.3960

YA(  3) =  0.6757
Y (  3) =  0.6980

YA(  4) =  0.7899
Y (  4) =  0.7880

YA(  5) =  0.8803
Y (  5) =  0.8630

YA(  6) =  1.0006
Y (  6) =  1.0000
```

UNCOMPLETED SUBSYSTEMS

STRUCTURE (PRIMARY)

THE SUM OF LEAST SQUARES USING THE I-TH COMPLETED SUBSYSTEM IS,

SLS(1) = 0.40044270E-03

SLS(2) = 0.97621518E-04

SLS(3) = 0.36122584E-02

SLS(4) = 0.33730622E-02

SLS(5) = 0.62553304E-03

SLS(6) = 0.26911236E-02

SLS(7) = 0.30279424E-03

THE CURVE WHICH BEST FITS THE DATA IS (2).

GUIDANCE AND NAVIGATION

THE NO. 1 INTERMEDIATE COST IS $ 16862.3

THE NO. 2 INTERMEDIATE COST IS $ 17381.9

THE RUN-OUT COST IS $ 17861.7

INSTRUMENTATION

THE SUM OF LEAST SQUARES USING THE I-TH COMPLETED SUBSYSTEM IS,

SLS(1) = 0.30832698E-04

SLS(2) = 0.10768677E-03

SLS(3) = 0.21780565E-02

SLS(4) = 0.16217203E-02

THE RUN-OUT COST IS $ 10136.6

REACTION CONTROL

HE SUM OF LEAST SQUARES USING THE I-TH COMPLETED SUBSYSTEM IS,

SLS(1) = 0.11010226E-02

SLS(2) = 0.17622352E-02

SLS(3) = 0.51921197E-03

SLS(4) = 0.29024102E-03

SLS(5) = 0.60860877E-03

SLS(6) = 0.22850492E-03

SLS(7) = 0.16982195E-02

THE CURVE WHICH BEST FITS THE DATA IS (6).

PROPULSION

THE NO. 1 INTERMEDIATE COST IS $ 22244.8

THE NO. 2 INTERMEDIATE COST IS $ 22250.6

THE RUN-OUT COST IS $ 21894.2

LAUNCH ESCAPE

HE SUM OF LEAST SQUARES USING THE I-TH COMPLETED SUBSYSTEM IS,

SLS(1) = 0.51448849E-03

SLS(2) = 0.10585605E-02

SLS(3) = 0.77737073E-03

SLS(4) = 0.26062358E-03

SLS(5) = 0.99991638E-04

SLS(6) = 0.11558883E-02

SLS(7) = 0.22257988E-04

THE CURVE WHICH BEST FITS THE DATA IS (7),
STRUCTURE (ADAPTER)

THE NO. 1 INTERMEDIATE COST IS $ 6023.5

THE NO. 2 INTERMEDIATE COST IS $ 6052.3

THE RUN-OUT COST IS $ 6022.0

COMMUNICATIONS

THE SUM OF LEAST SQUARES USING THE I-TH COMPLETED SUBSYSTEM IS,

SLS(1) = 0.10084036E-02

SLS(2) = 0.43504378E-03

SLS(3) = 0.47460056E-02

SLS(4) = 0.49065917E-02

SLS(5) = 0.12758655E-02

SLS(6) = 0.40859870E-02

SLS(7) = 0.94758913E-03

THE CURVE WHICH BEST FITS THE DATA IS (2),
GUIDANCE AND NAVIGATION

THE NO. 1 INTERMEDIATE COST IS $ 9569.5

THE NO. 2 INTERMEDIATE COST IS $ 9864.3

```
          SLS(  5) = 0.22353748E-03

          SLS(  6) = 0.98338142E-04

          SLS(  7) = 0.85920865E-03

  THE CURVE WHICH BEST FITS THE DATA IS (   6),
                 PROPULSION

NO. 1 INTERMEDIATE COST IS $      3723.0

NO. 2 INTERMEDIATE COST IS $      3723.9

      THE RUN-OUT COST IS $       3664.3
```